Sensitive and specific method for detecting G protein–coupled receptor mRNAs

Arne Hansen[1,3], Yidong Chen[2], Jason M Inman[5], Quang N Phan[4,6], Zhi-Qing Qi[4], Charlie C Xiang[4,6], Miklos Palkovits[4], Natasha Cherman[1], Sergei A Kuznetsov[1], Pamela G Robey[1], Eva Mezey[1,3] & Michael J Brownstein[4,6]

G protein–coupled receptors (GPCRs) mediate effects of extracellular signaling molecules in all the body's cells. These receptors are encoded by scarce mRNAs; therefore, detecting their transcripts with conventional microarrays is difficult. We present a method based on multiplex PCR and array detection of amplicons to assay GPCR gene expression with as little as 1 μg of total RNA, and using it, we profiled three human bone marrow stromal cell (BMSC) lines.

Cell signaling is mediated by many extracellular agents and receptors that respond to them, including the families of proteins that interact with intracellular signal cascades via trimeric G proteins. The 359 human GPCRs are involved in many physiological and pathological processes; and approximately 30% of all drugs in use today modulate GPCR activity[1,2]. Descriptions of many of the GPCRs are available online (http://www.iuphar-db.org/ and http://www.signaling-gateway.org/molecule/).

Despite their importance, our knowledge of the distributions and functions of GPCRs in cells and tissues of the body is incomplete. In part, this is because they are encoded by nonabundant mRNAs and are fully functional at levels $<1 \times 10^4$ protein molecules per cell[3]. It is difficult to detect GPCR transcripts in cell lines and even harder to do so in tissues, as the latter are comprised of many different cell types. Determining the expression profiles of GPCRs with conventional cDNA or oligonucleotide arrays has been very difficult to do. Worse yet, high sequence similarity of subfamily members results in 'cross-talk' in all but the best designed arrays.

We developed a new method to profile GPCR mRNAs in cell or tissue extracts. We used one-step reverse transcription PCR (RT-PCR) driven by eight sets of 50 primer pairs to amplify transcripts of interest (**Fig. 1a**). We pooled the amplicons from multiple reactions, dye-labeled them and hybridized them to an array of 55-base oligonucleotides corresponding to regions in 401 mRNAs of interest that were distinct from the primer sequences (**Fig. 1b**). Each gene is represented on our array by spotted sense and antisense oligonucleotides. Both are necessary because side reactions cannot be eliminated in multiplex RT-PCRs. These can lead to unequal consumption of primers that comprise each pair, unbalanced reactions, and generation of only the sense or antisense product.

Initially we used our assay to profile 100-ng samples of total RNA from three separate BMSC lines derived from a 45-year-old woman (BMSC1), a 9-year-old girl (BMSC2) and a 19-year-old man (BMSC3). Use of bone marrow from human subjects was approved by the Internal Review Board (National Institutes of Health protocol 94-D-0188). We assayed each sample of total RNA (that is, PCR-amplified) three times and developed three different arrays per amplicon each time. The results for individual samples were reproducible (**Fig. 2a–c** and **Supplementary Table 1** online). About 80% of the transcripts detected in any given template were found in all three PCR products. Among the rest, 12.7% were found in two out of three and 6.2% in only one assay. We used quantitative RT-PCR to measure the levels of transcripts that were not invariably detected in BMSC2 RNA. These mRNAs are indeed present (see **Supplementary Table 2** online).

An average of 85% of the receptor transcripts found in any given sample of BMSC total RNA, can be detected with a single set of PCRs (**Fig. 2**). For the most comprehensive profiles possible, we recommend amplifying samples two or three times to obtain 95% or virtually 100% of the signals our method can generate. A fourth replica yields few (1–3) additional species and is not cost effective (data not shown). Consequently, we listed all of the receptors detected in 1, 2 or 3 assays of BMSC1, 2 and 3, and used the lists to make a Venn diagram (**Fig. 2d**). The three BMSC samples had profiles that were quite similar. Each RNA sample had about 200 receptor transcripts, and the three individual samples had 170 mRNAs in common. Thus, the three BMSC lines we profiled seem remarkably similar given the fact that they were derived from different subjects, cultured for different amounts of time and grown in different media. The functional importance of those receptors that these lines do not share remains to be determined. A brief description of the receptors that we detected versus those reported in the literature is available in **Supplementary Table 1**.

We determined the false positive rate and sensitivity of our assay experimentally (**Supplementary Note 1** online). The former was 1–3%; the latter, about one transcript in a million.

[1]Craniofacial and Skeletal Diseases Branch, National Institute of Dental and Craniofacial Research, [2]Cancer Genetics Branch, National Human Genome Research Institute, [3]Laboratory of Neurochemistry, National Institute of Neurological Disorder and Stroke, [4]National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA. [5]The Institute for Genomic Research, Rockville, Maryland 20850, USA. [6]J. Craig Venter Institute, Rockville, Maryland 20850, USA. Correspondence should be addressed to A.H. (arnehansen@nidcr.nih.gov) or M.J.B. (mike@venterinstitute.org).

RECEIVED 10 APRIL; ACCEPTED 4 OCTOBER; PUBLISHED ONLINE 19 NOVEMBER 2006; DOI:10.1038/NMETH977

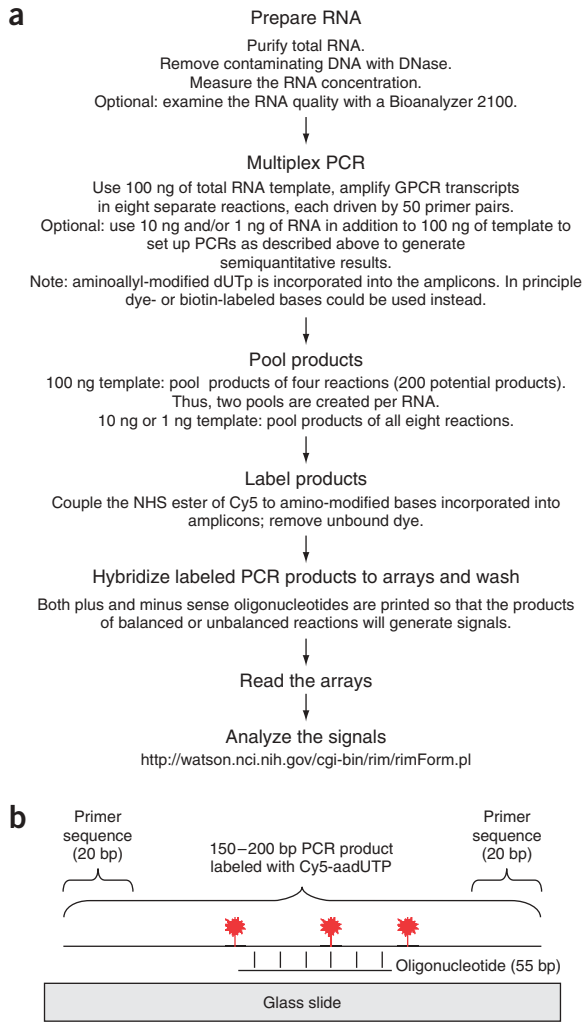

Figure 1 | GPCR transcript detection. (**a**) An outline of GPCR transcript profiling method. Samples of total RNA are subjected to one-step multiplex RT-PCR. The amplicons from multiple reactions are pooled, labeled and hybridized to arrays. The signals detected are analyzed using software that can be accessed via the internet. (**b**) Primer and probe design. Sense and antisense oligonucleotides are printed on glass slides. They bind Cy5-labeled (red) 150–200 bp PCR products, which are short enough to amplify efficiently, but long enough to be purified after the labeling reaction.

Because only about half of the GPCR transcripts represented on our arrays were detected in BMSCs, we examined 100-ng samples of total RNA from five additional tissues—liver, kidney, ovary, hippocampus and cerebral (temporal) cortex (see **Supplementary Table 3** online). Of the GPCR transcripts that we targeted, 84% were expressed in one or more of the six tissues (including BMSCs) that we studied. The primers and probes designed to detect the rest may have been synthesized poorly, may not work in a multiplex PCR setting or may represent receptor RNAs that are expressed in tissues other than those profiled.

To determine the coverage of our method compared to the use of conventional arrays, we compared the total number of GPCRs that we detected in the five tissues above with published results based on the use of conventional arrays[4]. On average, we found 2.5 times more GPCR transcripts using our method. Furthermore, 15% of the GPCR signals seen with conventional arrays were reported to be false positives[4] compared with our empirically determined false positive rate of $\leq$3%. This is probably explained by the improved signal strength of our method plus the two levels of specificity that are built into it: (i) PCR primers designed to amplify specific products and (ii) oligonucleotide probes designed to capture specific amplicons.

Although our method was designed to be qualitative versus quantitative, it is useful to know which receptor mRNAs are found in relatively high versus low amounts. We therefore wondered whether the abundance of GPCR mRNAs could be estimated by assaying serial dilutions of template using our method. We reasoned that relatively abundant transcripts might be detected in 10 ng or even 1 ng of total RNA, whereas the more rare species would be detected only in 100-ng samples. As expected, the less template we used, the fewer GPCR mRNAs we detected in all three BMSC samples studied. For example, we detected a total of 199, 85 and 49 receptor-coding transcripts in 100 ng, 10 ng and 1 ng of BMSC3 RNA, respectively (**Fig. 2e**). Our assay seems to be more error-prone when template is limiting. (This is true of global RNA amplification methods, too[5]). In spite of this, we believe that assaying serially diluted samples provides useful semi-quantitative information.

We tested the semiquatitative nature of the assay by performing a quantitative RT-PCR (Q-PCR) study of two groups of GPCR transcripts with two independent sets of primer pairs

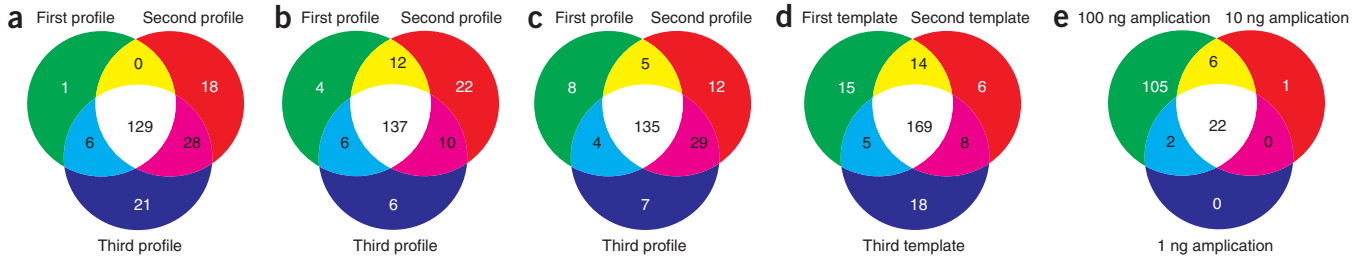

Figure 2 | Venn diagrams illustrating GPCR expression profiles of BMSCs. (**a**–**c**) The results obtained when BMSC1, 2 and 3, respectively, were amplified three different times. Three aliquots of each individual PCR product were applied to three separate microarrays, and mRNAs were called 'expressed' when signals were detected on all of the arrays. (**d**) A comparison of all receptors (detected in one, two or three PCR products) in each BMSC sample. Because detection of rare transcripts is a stochastic process and some transcripts will only be seen in the product of a single reaction, this is the most comprehensive profile we can obtain. The BMSC1, 2 and 3 receptor repertoires seem quite similar. (**e**) A comparison of receptor transcripts detected in one or more of three replica PCRs of 100, 10 and 1 ng of BMSC3 RNA. Most of the receptors detected in the 10 ng and 1 ng assays were found in the 100 ng assay.

(**Supplementary Note 2** online). The first group consisted of ten transcripts that were detected in 100 ng of RNA but not in 10 or 1 ng. The second group comprised 10 transcripts present in 100, 10 and 1 ng of RNA. There was a statistically significant difference in abundance between the members of these two groups in BMSC1 RNA (differences in mean cycle threshold values = 5.88 and 4.26 for the two sets of primers; $P = 5.76 \times 10^{-14}$ and 8.05×10^{-8}, respectively). Receptor transcripts that are consistently found in 1 ng, 10 ng and 100 ng of RNA appear to be more abundant than those that are only detected in 100 ng.

Thus, although signal strength in our assay does not reflect mRNA abundance, these results show that we could sort receptors into two abundance classes by assaying serial dilutions of template and excluding from consideration those transcripts that give inconsistent results. In general (but not always) the amounts of GPCR mRNAs and the amounts of proteins they encode are related[6]; furthermore, the concentration of agonist required to activate a GPCR is inversely related to the number of receptors on the plasma membrane. Regulation of GPCR activity is quite complex[7,8] however, and estimates of message abundance are no substitute for measures of signal transduction.

The method we have described is considerably more sensitive than conventional array techniques. It has allowed us to begin to determine which GPCR transcripts are ubiquitously expressed and which ones are unique to or abundant in a small number of human cells. Information of this sort should allow us and others to identify and validate GPCRs as targets for drug development. In fact, our receptor catalog has proven to be a useful starting point for studies of GPCR-mediated regulation of stem-cell division, apoptosis, differentiation, and migration *in vivo* and *in vitro*.

Detailed descriptions of the experimental procedures and data analysis are available **Supplementary Methods** and **Supplementary Note 3** online. The primer and probe sequences used are listed in **Supplementary Tables 4** and **5** online.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

This project was funded in part with Federal funds from the US National Cancer Institute, under Contract number N01-C0-12400, by the Intramural Research Programs of the National Institute of Mental Health, the National Institute of Neurological Disorder and Stroke, and the National Institute of Dental and Craniofacial Research, US National Institutes of Health, Deutsche Forschungsgemeinschaft Ha3423/2-1 and the J. Craig Venter Foundation.

DISCLAIMER

The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services, nor does mention of trade names, commercial products or organization imply endorsement by the US Government.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Drews, J. *Science* **287**, 1960–1964 (2000).
2. Wise, A., Gearing, K. & Rees, S. *Drug Discov. Today* **7**, 235–246 (2002).
3. Ostrom, R.S. & Insel, P.A. *Br. J. Pharmacol.* **143**, 235–245 (2004).
4. Hakak, Y., Shrestha, D., Goegel, M.C., Behan, D.P. & Chalmers, D.T. *FEBS Lett.* **550**, 11–17 (2003).
5. Xiang, C.C. *et al. Nucleic Acids Res.* **32**, e185 (2004).
6. Wynn, R.F. *et al. Blood* **104**, 2643–2645 (2004).
7. Sato, M., Blumer, J.B., Simon, V. & Lanier, S.M. *Annu. Rev. Pharmacol. Toxicol.* **46**, 151–187 (2006).
8. Luttrell, L.M. *Methods Mol. Biol.* **332**, 3–49 (2006).